Available online at www.sciencedirect.com
ScienceDirect
Ann. I. H. Poincaré – AN 26 (2009) 1081–1098
ANNALES DE L'INSTITUT HENRI POINCARÉ
ANALYSE NON LINÉAIRE
www.elsevier.com/locate/anihpc

Conjugate and cut loci of a two-sphere of revolution with application to optimal control

Lieu conjugué et lieu de coupure d'une 2-sphère de révolution et application en contrôle optimal

Bernard Bonnard [a,*], Jean-Baptiste Caillau [b], Robert Sinclair [c], Minoru Tanaka [d]

[a] *Institut de mathématiques de Bourgogne (UMR CNRS 5584), 9, avenue Savary, F-21078 Dijon, France*
[b] *ENSEEIHT-IRIT (UMR CNRS 5505), 2, rue Camichel, F-31071 Toulouse, France*
[c] *Mathematical Biology Unit, Okinawa Institute of Science and Technology, Okinawa Industrial Technology Center Annex, 12-2 Suzaki, Uruma, Okinawa 904-2234, Japan*
[d] *Department of Mathematics, Tokai University, Hiratsuka City, Kanagawa Pref., 259-1292, Japan*

Received 8 November 2007; received in revised form 13 February 2008; accepted 19 March 2008

Available online 21 June 2008

Abstract

The objective of this article is to present a sharp result to determine when the cut locus for a class of metrics on a two-sphere of revolution is reduced to a single branch. This work is motivated by optimal control problems in space and quantum dynamics and gives global optimal results in orbital transfer and for Lindblad equations in quantum control.

Résumé

Le but de cet article est de présenter une condition suffisante permettant de garantir que le lieu de coupure d'une classe de métriques sur la 2-sphère de révolution est réduit à une branche simple. Ce travail est motivé par des problèmes de contrôle optimal en mécanique spatiale et mécanique quantique. Des résultats globaux d'optimalité sont obtenus en transfert orbital ainsi que dans le cas des équations de Lindblad en contrôle quantique.

MSC: 53C20; 53C21; 49K15; 70Q05

Keywords: Conjugate and cut loci; 2-spheres of revolution; Optimal control; Space and quantum mechanics

* Corresponding author.
E-mail addresses: Bernard.Bonnard@u-bourgogne.fr (B. Bonnard), caillau@n7.fr (J.-B. Caillau), sinclair@oist.jp (R. Sinclair), m-tanaka@sm.u-tokai.ac.jp (M. Tanaka).

0294-1449/$ – see front matter
doi:10.1016/j.anihpc.2008.03.010

1. Introduction

The purpose of this article is to improve recent advanced results concerning the structure of the conjugate and cut loci on a two-surface of revolution [20,21] to analyze optimal control problems for both space and quantum control dynamics.

The determination of the cut and conjugate loci on a complete two-surface of revolution is a standard but difficult problem in Riemannian geometry. For a real analytic two-sphere, the cut locus at each point is a finite tree, whose extremities are conjugate points. This was stated by Poincaré [16] and proved by Myers [15]. Fleischmann [14] studied the behavior of geodesics on various surfaces of revolution in Euclidean space. We refer to [17] for modern tools of Riemannian manifolds and [19] for the behavior of geodesics on a surface of revolution.

The structure theorem of the cut locus of a point on a 2-dimensional Riemannian manifold was established by Hebda [12] and generalized by Shiohama and Tanaka [18] for the cut locus of a compact subset in an Alexandrov surface. By the structure theorem (see [18, Theorem A]), the cut locus is a local tree and is a union of countably many rectifiable Jordan arcs and the endpoints.

Still, precise computation is difficult and the complexity is in estimating the ramifying branches. Besides, a construction due to Gluck and Singer [10] proves that there exists a smooth strictly convex surface of revolution, homeomorphic to $\mathbb{S}^2$, whose cut locus is not stratifiable. The case of the triaxial ellipsoid has only recently been solved [13].

Even on an ellipsoid of revolution, the computation is not a standard exercise (in [3], the foreseen conjugate and cut loci were given as a conjecture). On an oblate ellipsoid the cut locus of a point different from the pole is a subarc of the antipodal parallel. For a prolate ellipsoid, the same holds replacing parallel by opposite half meridian. In the first case the Gaussian curvature is monotone increasing from the north pole to the equator and decreasing in the second case.

This result is a consequence of a general result in [21]: given a smooth metric on $\mathbb{S}^2$ of the form $dr^2 + m^2(r)\,d\theta^2$, where r is the angle along the meridian and θ the angle of revolution, assume the following:

1. $m(2a - r) = m(r)$ (reflective symmetry with respect to the equator, where $2a$ is the distance between poles).
2. The Gaussian curvature is monotone non-decreasing (resp. non-increasing) along a meridian from the north pole to the equator.

Then the cut locus of a point different from the pole is a simple branch located on the antipodal parallel (resp. opposite half meridian).

This gives a nice computable criterion to decide whether the cut locus is reduced to a simple branch. In parallel, in recent research projects on geometric optimal control in orbital transfer or quantum control, the optimality analysis can be reduced to an optimal control problem on a two-sphere of revolution for which the generalization of the previous result is crucial in several directions: first of all the monotonicity of the Gauss curvature is not satisfied, second the metric can have singularities. The key step is to relate the simple structure of the cut and conjugate loci to a tame property of the extremal flow.

The organization of this article is the following. In Section 2, we present the systems from space and quantum dynamics motivating the analysis. In Section 3, we give the sharp optimality result needed for the analysis in the Riemannian case, with application to our examples in Section 4. The analysis is extended in Section 5 to deal with almost Riemannian metrics on two-spheres of revolution encountered in our systems analysis and in extensions of the Gauss–Bonnet theorem [1]. In a concluding section we discuss in detail the contributions of this article as well as possible extensions.

It is also worth pointing out that our analysis is related to homotopy methods in optimal control, deforming the round sphere $\mathbb{S}^2$ and keeping simple conjugate and cut loci, to be compared to the opposite construction of [10] to generate complex such loci.

2. Motivating examples

2.1. Orbital transfer

The two-input coplanar transfer system [6] is modelled by a 2π-*periodic* system on a three-dimensional manifold M, of the form:

$$\frac{dq}{dl} = u_1 F_1(q,l) + u_2 F_2(q,l)$$

where the state q represents the geometric coordinates of the osculating ellipse, e.g. $q = (n, e, \theta)$, where n is the *mean motion*, e the *eccentricity* and θ the *argument of the pericenter*. The angular variable is the *longitude* $l \in \mathbb{S}^1$, while the trajectories are parameterized by the *cumulated* longitude $l \in \mathbb{R}$. When considering the energy minimization problem, the maximum principle tells us that minimizers have to be selected among extremal curve solutions of the Hamiltonian

$$H(q,p,l) = \frac{1}{2}\big(H_1^2 + H_2^2\big)(q,p,l)$$

where the H_i's are the Hamiltonian lifts $H_i(q,p,l) = \langle p, F_i(q,l)\rangle$, $i = 1, 2$. If low thrust is applied, we consider the long time behavior of the system which is approximated using the *averaged* Hamiltonian

$$\bar{H}(q,p) = \frac{1}{2\pi}\int_0^{2\pi} H(q,p,l)\,dl.$$

Averaging generates Lie brackets of the initial vector fields, so that the averaged Hamiltonian turns out to be a full rank quadratic form in the adjoint variable p, thus associated to a Riemannian metric. The computed expression is

$$\bar{g} = \frac{dn^2}{9n^{1/3}} + \frac{2n^{5/3}}{5(1-e^2)}\,de^2 + \frac{2n^{5/3}}{5-4e^2}e^2\,d\theta^2.$$

Such a metric can be normalized with $n = (5\rho/2)^{6/5}$, $e = \sin r$ so that:

$$\bar{g} = d\rho^2 + (\rho^2/c^2)g$$

where $c = \sqrt{2/5}$ and $g = dr^2 + m^2(r)\,d\theta^2$ with

$$m^2(r) = \frac{\sin^2 r}{1 - (4/5)\sin^2 r}.$$

By homogeneity, we can restrict our optimality analysis to the Riemannian metric g with $r \in [0, \pi/2]$, where the pole $e = 0$ corresponds to circular orbits, while $e = 1$ corresponds to parabolic orbits. It can be extended to an *analytic metric* on a two-sphere of revolution, where (r, θ) are the spherical coordinates.

If we now come back to the original system modelling coplanar orbital transfer and fix the direction of the control, we obtain a single-input periodic system of the form:

$$\frac{dq}{dl} = uF(q,l)$$

and an interesting case motivated by cone constraint due to electro-ionic propulsion is the so-called *tangential case* where the control has to be directed by velocity. Averaging, we obtain again a full rank Hamiltonian. A remarkable feature is that the same geometric coordinates remain orthogonal for the new metric [5]

$$\overline{g_t} = \frac{dn^2}{9n^{1/3}} + n^{5/3}\left[\frac{1+\sqrt{1-e^2}}{4(1-e^2)^{3/2}}\,de^2 + \frac{1+\sqrt{1-e^2}}{4(1-e^2)}e^2\,d\theta^2\right].$$

In the two-input case, the change of variables $e = \sin r$ only consisted in lifting the Poincaré disk on which (e, θ) are polar coordinates onto $\mathbb{S}^2$, where (r, θ) are the standard angles. To normalize now, we again set $n = (5\rho/2)^{6/5}$ and slightly twist the previous lifting according to

$$e = \sin r\sqrt{1 + \cos^2 r}$$

to obtain the normal form

$$\overline{g_t} = d\rho^2 + (\rho^2/c_t)g_t$$

with $c_t = c^2 = 2/5$ and $g_t = dr^2 + m_t^2(r)\,d\theta^2$:

$$m_t^2 = \sin^2 r\left(\frac{1 - (1/2)\sin^2 r}{1 - \sin^2 r}\right)^2.$$

In both cases we can define a homotopy deforming the round metric on $\mathbb{S}^2$, introducing $g_\lambda = dr^2 + XR_\lambda(X)\,d\theta^2$, setting $X = \sin^2 r$, while $R_0 = 1$, $R_\lambda(X) = R(\lambda X)$ and we get the

- *bi-input case* $R(X) = 1/(1-X)$ in which we have for $\lambda \in [0,1]$ a homotopy from the round metric $\lambda = 0$ to the orbital transfer for which $\lambda = 4/5$, while the limit case $\lambda = 1$ is *singular*, since R has a pole at $X = 1$;
- *tangential case* $R(X) = \{(1-X/2)/(1-X)\}^2$ in which we have for $\lambda \in [0,1]$ a homotopy from the round metric $\lambda = 0$ to the *orbital transfer which is singular*, since R has a pole of order two at $X = 1$.

2.2. Quantum control

We consider a dissipative two-level quantum system whose dynamics is governed by the *Lindblad equation*, see [22,7] for the details:

$$\begin{aligned}
\dot q_1 &= -\Gamma q_1 + u_2 q_3,\\
\dot q_2 &= -\Gamma q_2 - u_1 q_3,\\
\dot q_3 &= \gamma_- - \gamma_+ q_3 + (u_1 q_2 - u_2 q_1)
\end{aligned}$$

where the state space $q = (q_1, q_2, q_3)$ is restricted to the *Bloch ball*: $q_1^2 + q_2^2 + q_3^2 \leqslant 1$, and the control is of the form $u = u_1 + iu_2$, u_1, u_2 being two real functions, $|u| \leqslant 1$ and the three parameters Γ, γ_-, γ_+ describing the interaction with the environment and satisfying constraints: $\Gamma \geqslant \frac{\gamma_+}{2} \geqslant 0$, $\gamma_+ \geqslant |\gamma_-|$.

To minimize the effect of dissipation, we consider the problem of minimizing time of transfer, but the energy minimization problem shares similar properties.

The system is written

$$\dot q = F_0(q) + u_1 F_1(q) + u_2 F_2(q)$$

and introducing the Hamiltonian lifts $H_i = \langle p, F_i(q)\rangle$, $i = 0, 1, 2$, outside the switching surface $H_i = 0$, $i = 1, 2$, the maximal principle tells us that time optimal control trajectories are extremal solutions of the Hamiltonian vector field:

$$H(q,p) = H_0(q,p) + \left(\sum_{i=1}^{2} H_i^2(q,p)\right)^{1/2}.$$

Since $|u| \leqslant 1$, the problem is invariant by change of coordinates and feedback transformations of the form $u = \beta(q)v$, where $\beta(q)$ is an orthogonal matrix.

We consider only the case where $\gamma_- = 0$. Since the Bloch ball is invariant, we introduce the spherical coordinates $q_3 = r\cos\phi$, $q_1 = r\sin\phi\cos\theta$, $q_2 = r\sin\phi\sin\theta$, using the relations

$$\begin{aligned}
r^2 &= q_1^2 + q_2^2 + q_3^2,\\
\rho &= \ln r,\\
\theta &= \arctan q_2/q_1,\\
\phi &= \arccos q_3/r
\end{aligned}$$

and the feedback transformation $v_1 = \cos\theta u_1 + u_2\sin\theta$, $v_2 = -\sin\theta u_1 + u_2\cos\theta$. The system then takes the form

$$\begin{aligned}
\dot\rho &= -(\Gamma\sin^2\phi + \gamma_+\cos^2\phi),\\
\dot\theta &= -(\operatorname{cotan}\phi)v_1,\\
\dot\phi &= v_2 + \frac{\sin 2\phi}{2}(\gamma_+ - \Gamma).
\end{aligned}$$

If we consider the Hamiltonian system describing the evolution of generic extremals we get

$$H = -p_\rho(\Gamma\sin^2\phi + \gamma_+\cos^2\phi) + p_\phi\frac{\sin 2\phi}{2}(\gamma_+ - \Gamma) + R,$$

where R is respectively $(H_1^2 + H_2^2)^{1/2}$ in the time minimization case and $\frac{1}{2}(H_1^2 + H_2^2)$ in the energy minimization case, with:

$$H_1 = -p_\theta \operatorname{cotan}\phi,$$
$$H_2 = p_\phi.$$

We observe the following:

- *If* $\gamma_+ = \Gamma$, both extremal flows are a suspension of the extremal flows associated to the metric on the two-sphere $\mathbb{S}^2$ with coordinates $(r = \phi, \theta)$ given by

$$g = dr^2 + (\tan^2 r)\, d\theta^2,$$

  which corresponds to the limit case $\lambda = 1$ for the homotopy of the previous section, in the bi-input case.
- *If* $\gamma_+ \neq \Gamma$, the Hamiltonians admit ρ and θ as cyclic coordinates, hence p_ρ and p_θ are first integrals. A deeper analysis reveals that the optimality analysis is related to a geometric perturbation of the previous case for which:
  - For fixed p_ρ, the reduced Hamiltonians are associated with a one parameter family of optimal control problems on the two-sphere of revolution.
  - For each such problem, the extremal flow shares similar properties to the case $\gamma_+ = \Gamma$, which makes the computation of the conjugate and cut loci tractable.

3. Conjugate and cut loci on a two-sphere of revolution

The objective of this section is to characterize when the cut locus of a point different from the pole on a two-sphere of revolution is reduced to a single segment and the conjugate locus has the standard astroid shape. This is based mainly on the analysis in [21] but extensions are decoded from the properties of the extremal flow only. This is not restrictive since in complete Riemannian 2D-manifolds the computation of the cut locus is obtained by evaluating the separating line of a point q_0, $L(q_0)$ where minimizers starting from q_0 are intersecting while the conjugate locus is the set of limit points of intersecting neighboring extremals or equivalently the envelope of such extremals.

A compact Riemannian manifold (M, g) homeomorphic to a 2-sphere is called a 2-*sphere of revolution*, if M admits a point p such that for any two points q_1, q_2 on M with $d(p, q_1) = d(p, q_2)$, where $d(\cdot,\cdot)$ denotes the Riemannian distance function, there exists an isometry f on M satisfying $f(q_1) = q_2$ and $f(p) = p$. The point p is called a *pole* of M. It is proved in [21] that each pole of a 2-sphere of revolution has a unique cut point, which is also a pole of the 2-sphere. By fixing a pole p on a 2-sphere M of revolution, we introduce geodesic polar coordinates (r, θ) around the pole p. The Riemannian metric g is expressed as $g = dr^2 + m(r)^2\, d\theta^2$ on $M \setminus \{p, q\}$, where

$$m\big(r(x)\big) := \sqrt{g\left(\left(\frac{\partial}{\partial\theta}\right)_x, \left(\frac{\partial}{\partial\theta}\right)_x\right)}, \tag{1}$$

and q denotes the unique cut point of p. The Gaussian curvature G at a point $x \in M \setminus \{p, q\}$ is equal to

$$G(x) = -\frac{m''(r(x))}{m(r(x))}. \tag{2}$$

Each unit speed geodesic $\mu : \mathbb{R} \to M$ passing through the pole p is called a *meridian*. Since q is the unique cut point of p, μ passes through q. It is easily checked that μ is periodic, i.e., $\mu(r + 4a) = \mu(r)$, where $2a := d(p, q)$. Each curve $r = c \in (0, 2a)$ is called a *parallel*.

Let $\gamma(s) = (r(s), \theta(s))$ be a unit speed geodesic on the manifold M. Then, there exists a constant ν such that

$$m\big(r(s)\big)^2 \cdot \theta'(s) = m\big(r(s)\big) \cos\eta(s) = \nu \tag{3}$$

holds for any s, where $\eta(s)$ denotes the angle $\angle(\dot\gamma(s), (\partial/\partial\theta)_{\gamma(s)})$ made by $\dot\gamma(s) := d\gamma(\partial/\partial s)$ and $(\partial/\partial\theta)_{\gamma(s)}$. The relation (3) is called the *Clairaut relation*, and the constant ν is called the *Clairaut constant of* γ. Since γ is unit speed, it follows from (3) that

$$r'(s) = \varepsilon\big(r'(s)\big) \frac{\sqrt{m(r(s))^2 - \nu^2}}{m(r(s))}, \tag{4}$$

where $\varepsilon(r'(s))$ denotes the sign of $r'(s)$. In particular, $r'(s)=0$ if and only if $m(r(s))=|\nu|$. Hence, the geodesic γ stays in the closure of a connected component of $(m\circ r)^{-1}(|\nu|,\infty)$, and if $m(r(s))=|\nu|$ at $s=s_0$, then γ is tangent to the parallel $r=r(s_0)$. It follows from (3) and (4) that

$$\theta(s_2)-\theta(s_1)\equiv\varepsilon\bigl(r'(s)\bigr)\int\limits_{r(s_1)}^{r(s_2)} f(r,\nu)\,dr \mod 2\pi \tag{5}$$

holds, where

$$f(r,\nu)=\frac{\nu}{m(r)\sqrt{m(r)^2-\nu^2}},$$

if $r'(s)\neq 0$ on (s_1,s_2), and moreover the length $L(\gamma|_{[s_1,s_2]})$ of $\gamma|_{[s_1,s_2]}$ equals

$$L(\gamma|_{[s_1,s_2]})=\varepsilon\bigl(r'(s)\bigr)\int\limits_{r(s_1)}^{r(s_2)}\frac{m(r)}{\sqrt{m(r)^2-\nu^2}}\,dr \tag{6}$$

if $r'(s)\neq 0$ on (s_1,s_2). Hereafter, *we assume that the function m satisfies*

$$m(r)=m(2a-r) \tag{7}$$

for any $r\in(0,2a)$, where $2a=d(p,q)$. The parallel $r=a$ is called the *equator* of M. By (7), M is reflectively symmetric with respect to the equator.

For technical reasons, we introduce the Riemannian universal covering manifold

$$\widetilde{M}:=\bigl((0,2a)\times\mathbb{R},d\tilde r^2+m(\tilde r)^2\,d\tilde\theta^2\bigr)$$

of $(M\setminus\{p,q\},dr^2+m(r)^2\,d\theta^2)$. Note that Eqs. (3), (4), and (6) hold for geodesics on $\widetilde{M}$. Eq. (5) is replaced by

$$\tilde\theta\bigl(\tilde\gamma(s_2)\bigr)-\tilde\theta\bigl(\tilde\gamma(s_1)\bigr)=\varepsilon\bigl((\tilde r\circ\tilde\gamma)'(s)\bigr)\int\limits_{\tilde r(\tilde\gamma(s_1))}^{\tilde r(\tilde\gamma(s_2))} f(r,\nu)\,dr. \tag{8}$$

Here, we assume that $(\tilde r\circ\tilde\gamma)'(s)\neq 0$ on (s_1,s_2). For each $\nu\in[0,m(a)]$, let $\tilde\gamma_\nu$ denote a unit speed geodesic on $\widetilde{M}$ with the Clairaut constant ν emanating from a point on $\tilde r^{-1}(a)$. Since $\tilde\gamma_\nu$ satisfies the Clairaut relation,

$$\angle\bigl(\dot{\tilde\gamma}_\nu(0),(\partial/\partial\tilde\theta)_{\tilde\gamma_\nu(0)}\bigr)=\arccos\frac{\nu}{m(a)}$$

holds.

Lemma 3.1. *If* $m'\neq 0$ *on* $(0,a)$, *then for each* $\nu\in(0,m(a))$, *the geodesic* $\tilde\gamma_\nu$ *intersects* $\tilde r=a$ *again at a point* $\tilde\gamma_\nu(t_0(\nu))$, *and the function, which is called the half period function of* $\widetilde{M}$,

$$\varphi(\nu):=2\int\limits_{\xi(\nu)}^{a}\frac{\nu}{m(r)\sqrt{m(r)^2-\nu^2}}\,dr \tag{9}$$

is well-defined and is equal to $\tilde\theta(\tilde\gamma_\nu(t_0(\nu)))-\tilde\theta(\tilde\gamma_\nu(0))$. *Here,* $\xi(\nu):=(m|_{[0,a]})^{-1}(\nu)$.

Proof. Choose any $\nu\in(0,m(a))$ and fix it. We may assume that $(\tilde r\circ\tilde\gamma_\nu)'(0)<0$, since (7) holds. It is clear from (4) that $\tilde\gamma_\nu$ stays in $\tilde r^{-1}[\xi(\nu),2a-\xi(\nu)]$. Since $m'\neq 0$ on $(0,a)$, it follows from [19, Lemma 7.1.7] that $(\tilde r\circ\tilde\gamma_\nu)'(t_1)=0$ for some $t_1>0$, i.e., $\tilde\gamma_\nu$ is tangent to the parallel arc $\tilde r=\xi(\nu)$ at $\tilde\gamma_\nu(t_1)$. From (8), we get

$$\tilde\theta\bigl(\tilde\gamma_\nu(t_1)\bigr)-\tilde\theta\bigl(\tilde\gamma_\nu(0)\bigr)=\frac12\varphi(\nu). \tag{10}$$

Since $(\tilde r\circ\tilde\gamma_\nu)'(t)>0$ for any $t>t_1$ sufficiently close to t_1 and $(\tilde r\circ\tilde\gamma_\nu)'(t)\neq 0$ on (t_1,t) if $\tilde r\circ\tilde\gamma_\nu<2a-\xi(\nu)$, there exists $t_0(\nu)$ $(>t_1)$ such that $\tilde r\circ\tilde\gamma_\nu(t_0(\nu))=a$ and $\tilde r\circ\tilde\gamma_\nu<a$ on $(t_1,t_0(\nu))$. Hence $\tilde\gamma_\nu$ intersects $\tilde r=a$ again at the point $\tilde\gamma_\nu(t_0(\nu))$. Since $\tilde\theta(\tilde\gamma_\nu(t_0(\nu)))-\tilde\theta(\tilde\gamma_\nu(t_1))$ is equal to $\varphi(\nu)/2$, the proof of our lemma is complete. $\square$

It is not difficult to calculate the parameter value $t_0(\nu)$ in Lemma 3.1. From (6) and (7), we get

$$t_0(\nu)=2\int_{\xi(\nu)}^{a}\frac{m(r)}{\sqrt{m(r)^2-\nu^2}}\,dr. \tag{11}$$

Since

$$\frac{m}{\sqrt{m^2-\nu^2}}=\frac{\sqrt{m^2-\nu^2}}{m}+\frac{\nu^2}{m\sqrt{m^2-\nu^2}},$$

we obtain

$$t_0(\nu)=2\int_{\xi(\nu)}^{a}\frac{\sqrt{m(r)^2-\nu^2}}{m(r)}\,dr+\nu\varphi(\nu). \tag{12}$$

Lemma 3.2. *If the cut locus of a point q on the equator $r=a$ is a subset of the equator, then the function $\varphi(\nu)$ is well-defined and monotone non-increasing on $(0,m(a))$.*

Proof. If the cut locus C_q of q consists of a single point, then it is clear that $\varphi(\nu)$ is constant. Thus, we may assume that C_q is a subarc of the equator. Let q_0 be an endpoint of C_q. First we will prove that $q_0:=\gamma_{m(a)}(t_0)$ is conjugate to $q:=\gamma_{m(a)}(0)$ along $\gamma_{m(a)}$, where $\gamma_{m(a)}$ denotes the subarc of the equator joining q to q_0. Since $\gamma_{m(a)}$ does not contain a cut point of q in its interior, it is a minimal geodesic segment joining q to q_0. If $\gamma_{m(a)}$ is the unique minimal geodesic segment joining q to q_0, then it is clear that q_0 is conjugate to q along $\gamma_{m(a)}$. Hence we suppose that there exists a minimal geodesic segment $\alpha:[0,t_0]\to M$ joining q and q_0 which bounds a disc domain D together with $\gamma_{m(a)}|_{[0,t_0]}$. Here, we may assume that $(r\circ\alpha)'(0)<0$ and $(r\circ\gamma_{m(a)})'(0)<0$ by (7). Since D has no cut point of q, any geodesic segment β emanating from q with $(r\circ\beta)'(0)<0$ must pass through the point q_0, if $\angle(\dot\beta(0),\dot\gamma_{m(a)}(0))<\angle(\dot\alpha(0),\dot\gamma_{m(a)}(0))$. Thus, we get a geodesic variation of $\gamma_{m(a)}|_{[0,t_0]}$, which is a family of geodesic segments joining q to q_0. Hence, we have proved that q_0 is a conjugate point of q along $\gamma_{m(a)}$. This implies that the Gaussian curvature G is positive on the equator. Since $m''(a)=-G(q)m(a)<0$ by (2) and $m'(a)=0$, m' is positive on $(a-\delta,a)$ for some $\delta>0$. Suppose that $m'(b)=0$ for some $b\in(0,a)$. From [19, Lemma 7.1.4], the parallel $r=b$ is a geodesic. By choosing the maximal $b(<a)$ satisfying $m'(b)=0$, we may assume that m' is positive on (b,a). Thus, $m(r)>m(b)$ on $(b,a]$. Suppose that the geodesic $\gamma_{m(b)}$ is tangent to a parallel. Here, for each $\nu\in[0,m(a))$, γ_ν denotes the unit speed geodesic emanating from q with the Clairaut constant ν satisfying $(r\circ\gamma_\nu)'(0)<0$. From (4), the possible parallel, to which $\gamma_{m(b)}$ is tangent, is the geodesic parallel $r=b$, but $\gamma_{m(b)}$ cannot be tangent to another geodesic $r=b$. Hence, $(r\circ\gamma_{m(b)})'(0)\neq 0$, and in particular, $\gamma_{m(b)}$ does not intersect the equator again (see [19, Fig. 7.1.2] on the behavior of such a geodesic). Since M is compact, there exists a cut point of q along $\gamma_{m(b)}$. This contradicts the assumption that C_q is a subarc of the equator. Therefore, m' is non-zero on $(0,a)$ and the function $\varphi(\nu)$ is well-defined on $(0,m(a))$. It is now clear that γ_ν intersects the equator again at $\gamma_\nu(t_0(\nu))$ and $\theta(\gamma_\nu(t))\leqslant\pi$ for any $\nu\in(0,m(a))$ and any $t\in(0,t_0(\nu)]$. Here, for a technical reason, the geodesic polar coordinates (r,θ) are chosen so as to satisfy $\theta(\gamma_\nu(0))=\theta(q)=0$. In particular, $\varphi(\nu)\leqslant\pi$ for any $\nu\in(0,m(a))$.

Next, we will prove that φ is monotone non-increasing. Choose any two numbers $\nu_1<\nu_2$ in $(0,m(a))$ and fix them. By the Clairaut relation and the inequality,

$$\arccos\frac{\nu_1}{m(a)}>\arccos\frac{\nu_2}{m(a)},$$

the geodesic $\gamma_{\nu_2}|_{(0,\delta)}$ lies in the domain D_{ν_1} bounded by the equator and $\gamma_{\nu_1}|_{[0,t_0(\nu_1)]}$ for some $\delta>0$. Since C_q is a subset of the equator, the geodesic segment γ_{ν_2} lying in D_{ν_1} does not pass through $\gamma_{\nu_1}|_{(0,t_0(\nu_1))}$, but passes through the equator and intersects at $\gamma_{\nu_2}(t_0(\nu_2))$. Thus, $\gamma_{\nu_2}(t_0(\nu_2))$ is a point on the subarc of the equator with endpoints $\gamma_{\nu_1}(0)$ and $\gamma_{\nu_1}(t_0(\nu_1))$, and in particular, $\varphi(\nu_2)\leqslant\varphi(\nu_1)$ holds. □

Lemma 3.3. *If $m'\neq 0$ on $(0,a)$ and the function $\varphi:(0,m(a))\to\mathbb{R}$ is monotone non-increasing, then the cut locus of each point on the equator $r=a$ is a subset of the equator.*

Proof. Let γ_ν, $\nu \in [0, m(a)]$, denote the geodesic emanating from a point q on the equator that was defined in the proof of Lemma 3.2. It follows from Lemma 3.1 that for each $\nu \in [0, m(a))$, γ_ν intersects the equator again at $\gamma_\nu(t_0(\nu))$. Choose any $\nu \in (0, m(a))$ and fix it. It follows from [19, Proposition 7.2.3] that $\gamma_\nu(t_1)$ is the first conjugate point of q along γ_ν if and only if

$$\frac{\partial\theta}{\partial\nu}\bigl(r\bigl(\gamma_\nu(t_1)\bigr), \nu\bigr) = 0, \tag{13}$$

where

$$\theta(r, \nu) := \int_{\xi(\nu)}^{a} f(r, \nu)\,dr + \int_{\xi(\nu)}^{r} f(r, \nu)\,dr. \tag{14}$$

Notice that it follows from [19, Proposition 7.2.2] and [19, Corollary 7.2.1] that there is no conjugate point of q along $\gamma_\nu|_{[0,t]}$, if $(r \circ \gamma_\nu)' \neq 0$ on $[0, t)$. It is clear from (7) that

$$\theta(r, \nu) = \varphi(\nu) - \int_{r}^{a} f(r, \nu)\,dr \tag{15}$$

holds. Hence,

$$\frac{\partial\theta}{\partial\nu}\bigl(r \circ \gamma_\nu(t), \nu\bigr) = \varphi'(\nu) - \int_{r(\gamma_\nu(t))}^{a} f_\nu(r, \nu)\,dr < 0, \tag{16}$$

if $r \circ \gamma_\nu(t) < a$, since $\varphi'(\nu) \leqslant 0$ on $(0, m(a))$. Note that

$$f_\nu(r, \nu) = \frac{m(r)}{(m(r)^2 - \nu^2)^{3/2}} > 0.$$

Thus, there is no conjugate point of q along $\gamma_\nu|_{[0,t]}$, if $\gamma_\nu([0, t]) \subset r^{-1}(0, a)$. By taking the limit in (16), we may prove that there is no conjugate point of q along $\gamma_{m(0)}|_{[0,t]}$, if $\gamma_{m(0)}([0, t]) \subset r^{-1}[0, a)$. Therefore, we have proved that there is no conjugate point of q along $\gamma_\nu|_{[0,t]}$, if $\gamma_\nu([0, t]) \subset r^{-1}[0, a)$ and $\nu \in [0, m(a))$.

Suppose that there exists a cut point $x \notin r^{-1}(a)$ of q. From (7), we may assume that x is a point in $r^{-1}[0, a)$. Let $\gamma_{\nu_1} : [0, d(q, x)] \to M$ denote a minimal geodesic segment joining q to x. We may assume that x is conjugate to q along γ_{ν_1}. Otherwise, there exists a minimal geodesic segment $\gamma_{\nu_1}|_{[0,d(q,x)]}$, $\nu_2 \in [0, m(a)) \setminus \{\nu_1\}$, joining q to x. Thus, both geodesic segments bound a disc domain D. Since the cut locus C_q in D has no circle, we may find an endpoint $y \in r^{-1}(0, a)$ in $C_q \cap D$. The endpoint y is conjugate to q along any minimal geodesic segment joining q to y. Therefore, by exchanging x and y, we may assume that x is conjugate to q along γ_{ν_1}. This contradicts the fact that there is no conjugate point of q along $\gamma_\nu|_{[0,t]}$, if $\gamma_\nu([0, t]) \subset r^{-1}[0, a)$. Therefore, the cut locus of q is a subset of the equator. □

Choose any point q in M, which is not a pole. We introduce geodesic polar coordinates (r, θ) around the pole p on M satisfying $\theta(q) = 0$. Put $u := r(q) \in (0, 2a)$. For each $\nu \in (0, m(u)]$, let $\alpha_\nu, \beta_\nu : [0, \infty) \to M$ denote the geodesics emanating from $q = \alpha_\nu(0) = \beta_\nu(0)$ with the Clairaut constant ν satisfying

$$(r \circ \alpha_\nu)'(0) \leqslant 0 \leqslant (r \circ \beta_\nu)'(0).$$

Hence, from the Clairaut relation,

$$\angle\bigl(\dot\alpha_\nu(0), (\partial/\partial\theta)_q\bigr) = \angle\bigl(\dot\beta_\nu(0), (\partial/\partial\theta)_q\bigr) = \arccos\frac{\nu}{m(u)}.$$

From (7), the geodesics α_ν and β_ν intersect again at $(r, \theta)^{-1}(2a - u, \varphi(\nu)) = \alpha_\nu(t_0(\nu)) = \beta_\nu(t_0(\nu))$ (see [21], or [11]).

Lemma 3.4. *Assume that $m' \neq 0$ on $(0, a)$ and that φ is monotone non-increasing. Then, for each $\nu \in (0, m(u)]$, $\alpha_\nu|_{[0,t_0(\nu)]}$ is minimal. Furthermore, each point of $r^{-1}(2a - u) \cap \theta^{-1}[\varphi(m(u)), \pi]$ is a cut point of q.*

Proof. Since the geodesic $\gamma_{m(0)}$ meets the equator again at the antipodal point of $\gamma_{m(0)}(0)$, we have $\lim_{\nu\downarrow 0}\varphi(\nu)=\pi$. Thus, $0<\varphi(\nu)\leqslant\pi$ for any $\nu\in(0,m(a))$, since φ is monotone non-increasing. If $\varphi(m(u))=\pi$, then $\varphi(\nu)=\pi$ for any $\nu\in[0,m(u)]$. Hence the cut locus of q consists of a single point. It is now clear that both geodesics $\alpha|_{[0,t_0(\nu)]}$ and $\beta|_{[0,t_0(\nu)]}$ are minimal for any $\nu\in[0,m(u)]$. Hence we may assume that $\varphi(m(u))<\pi$. Choose any point $x\in r^{-1}(2a-u)\cap\theta^{-1}[\varphi(m(u)),\pi)$. From (6), the length t_1 of α equals

$$\int_u^{2a-\xi(\nu_1)}\frac{m(r)}{\sqrt{m(r)^2-\nu_1^2}}\,dr+\int_{2a-u}^{2a-\xi(\nu_1)}\frac{m(r)}{\sqrt{m(r)^2-\nu_1^2}}\,dr \tag{17}$$

(if $\alpha=\beta_{\nu_1}|_{[0,t_1]}$), or to

$$\int_{\xi(\nu_1)}^{u}\frac{m(r)}{\sqrt{m(r)^2-\nu_1^2}}\,dr+\int_{\xi(\nu_1)}^{2a-u}\frac{m(r)}{\sqrt{m(r)^2-\nu_1^2}}\,dr \tag{18}$$

(if $\alpha=\alpha_{\nu_1}|_{[0,t_1]}$). By (7) and (11), both Eqs. (17) and (18) equal

$$t_1=2\int_{\xi(\nu_1)}^{a}\frac{m(r)}{\sqrt{m(r)^2-\nu^2}}\,dr=t_0(\nu_1). \tag{19}$$

Therefore, $\alpha_{\nu_1}|_{[0,t_0(\nu_1)]}$ and $\beta_{\nu_1}|_{[0,t_0(\nu_1)]}$ are minimal geodesic segments joining q to x. Furthermore, for any $\nu\in\varphi^{-1}(\theta(x))$, $\alpha_\nu|_{[0,t_0(\nu)]}$ and $\beta_\nu|_{[0,t_0(\nu)]}$ are minimal, since $t_0'(\nu)=\nu\varphi'(\nu)$ by (12). Since x is arbitrarily taken, this implies that $\alpha_\nu|_{[0,t_0(\nu)]}$ and $\beta_\nu|_{[0,t_0(\nu)]}$ are minimal for any $\nu\in(0,m(u)]$ with $\varphi(\nu)<\pi$, hence for any $\nu\in(0,m(u)]$ from the limit argument. Therefore, $\alpha_\nu|_{[0,t_0(\nu)]}$ is minimal for all $\nu\in(0,m(u))$. The second claim is clear, since each point of $r^{-1}(2a-u)\cap\theta^{-1}(\varphi(m(u)),\pi]$ is joined by two minimal geodesic segments $\alpha_\nu|_{[0,t_0(\nu)]}$ and $\beta_\nu|_{[0,t_0(\nu)]}$ for some $\nu\in[0,m(u))$, and the cut locus is closed. □

Theorem 3.5. *Let $(M,dr^2+m(r)^2\,d\theta^2)$ denote a 2-sphere of revolution, where $m:(0,2a)\to(0,\infty)$ is a smooth function satisfying* (7). *If the cut locus of a point on $r=a$ is a subset of $r=a$, then, the cut locus of a point q with $r(q)\in(0,2a)\setminus\{a\}$ is a subset of the antipodal parallel $r=2a-r(q)$.*

Proof. Let $q\in r^{-1}((0,2a)\setminus\{a\})$ be any point, and set $u:=r(q)$. Since M is reflectively symmetric with respect to the meridian passing through q, the set $r^{-1}(2a-u)\cap\theta^{-1}[\varphi(m(u)),2\pi-\varphi(m(u))]$ is a subset of C_q by Lemma 3.4. Choose any cut point x of q. Then, we have a minimal geodesic segment γ joining q to x. Since we may assume that $0<\theta(x)\leqslant\pi$, γ is equal to α_{ν_1}, or β_{ν_1} for some $\nu_1\in[0,m(u)]$. Here ν_1 denotes the Clairaut constant of γ. The point x is not an interior point of $\alpha_{\nu_1}|_{[0,t_0(\nu_1)]}$, or $\beta_{\nu_1}|_{[0,t_0(\nu_1)]}$, since both segments are minimal. Furthermore, γ is a subarc of $\alpha_{\nu_1}|_{[0,t_0(\nu_1)]}$, or $\beta_{\nu_1}|_{[0,t_0(\nu_1)]}$, since γ is minimal. Hence,

$$x=\alpha_{\nu_1}\bigl(t_0(\nu_1)\bigr)\quad\bigl(=\beta_{\nu_1}\bigl(t_0(\nu_1)\bigr)\bigr).$$

Since $r(\alpha_{\nu_1}(t_0(\nu_1)))=2a-u$ and $\varphi(m(u))\leqslant\varphi(\nu_1)=\theta(x)\leqslant\pi$, any cut point of q is a point of $r^{-1}(2a-u)\cap\theta^{-1}[\varphi(m(u)),2\pi-\varphi(m(u))]$, which is a subarc of the antipodal parallel of q. □

Theorem 3.6. *Let $(M,dr^2+m(r)^2\,d\theta^2)$ denote a 2-sphere of revolution, where $m:(0,2a)\to(0,\infty)$ is a smooth function satisfying* (7). *Assume that the cut locus of a point on $r=a$ is a subset of $r=a$. If the half period function φ defined by* (9) *is such that $\varphi''(\nu)\geqslant 0$ on $(0,m(a))$ and $\varphi''(\nu)>0$ whenever $\varphi'(\nu)=0$, then the first conjugate locus of any point q which is not a pole of M has exactly four cusps.*

Proof. Choose any point $q\in M$ which is not a pole. Set $u:=r(q)\in(0,2a)$. It follows from [19, Proposition 7.2.3] that $\alpha_\nu(t_c(\nu))$ is the first conjugate point of q along α_ν if and only if

$$\frac{\partial\theta_\alpha}{\partial\nu}\bigl(r\bigl(\alpha_\nu\bigl(t_c(\nu)\bigr)\bigr),\nu\bigr)=0, \tag{20}$$

where

$$\theta_\alpha(r,\nu)=\int_{\xi(\nu)}^{u} f(r,\nu)\,dr+\int_{\xi(\nu)}^{r} f(r,\nu)\,dr. \tag{21}$$

From (7), it is clear that

$$\theta_\alpha(r,\nu)=\varphi(\nu)-\int_{r}^{2a-u} f(r,\nu)\,dr \tag{22}$$

holds. Hence,

$$\frac{\partial\theta_\alpha}{\partial\nu}(r,\nu)=\varphi'(\nu)-\int_{r}^{2a-u} f_\nu(r,\nu)\,dr, \tag{23}$$

where

$$f_\nu(r,\nu)=\frac{m(r)}{(m(r)^2-\nu^2)^{\frac{3}{2}}}.$$

From (20) and (23), it follows that

$$\varphi'(\nu)=\int_{u_\alpha(\nu)}^{2a-u} f_\nu(r,\nu)\,dr, \tag{24}$$

where $u_\alpha(\nu)=r(\alpha_\nu(t_c(\nu)))$. Hence, the first conjugate point of q along α_ν is given by

$$\theta=\theta_\alpha\big(u_\alpha(\nu),\nu\big),\quad r=u_\alpha(\nu). \tag{25}$$

Since we assume that $\varphi'(\nu)\leqslant 0$ for each $\nu\in(0,m(u))$,

$$2a-u\leqslant u_\alpha(\nu) \tag{26}$$

by (24). Furthermore, $\varphi'(\nu)=0$ if and only if $2a-u=u_\alpha(\nu)$. By differentiating (24) with respect to ν, we have

$$\varphi''(\nu)+\int_{2a-u}^{u_\alpha(\nu)} f_{\nu\nu}(r,\nu)\,dr=-f_\nu\big(u_\alpha(\nu),\nu\big)u'_\alpha(\nu). \tag{27}$$

Since $f_{\nu\nu}(r,\nu)=3\nu m(r)(m(r)^2-\nu^2)^{-5/2}>0$ and $\varphi''(\nu)\geqslant 0$, $u'_\alpha(\nu)\leqslant 0$ on $(0,m(u))$. If $u'_\alpha(\nu)=0$, then, by (27), $\varphi''(\nu)=0$ and $2a-u=u_\alpha(\nu)$. This implies that $\varphi''(\nu)=\varphi'(\nu)=0$. This contradicts the assumption of our theorem. Therefore, $u'_\alpha(\nu)<0$ on $(0,m(u))$. In particular, there is no cusp on the open arc defined by (25), $\nu\in(0,m(u))$. Let $\nu\in(0,m(u))$ be any fixed number. The velocity vector $v_\alpha(\nu)$ of the curve defined by (25) is given by

$$v_\alpha(\nu)=f\big(u_\alpha(\nu),\nu\big)u'_\alpha(\nu)\left(\frac{\partial}{\partial\theta}\right)_{\alpha_\nu(t_c(\nu))}+u'_\alpha(\nu)\left(\frac{\partial}{\partial r}\right)_{\alpha_\nu(t_c(\nu))}. \tag{28}$$

Hence, $v_\alpha(\nu)$ is parallel to

$$f\big(u_\alpha(\nu),\nu\big)\left(\frac{\partial}{\partial\theta}\right)_{\alpha_\nu(t_c(\nu))}+\left(\frac{\partial}{\partial r}\right)_{\alpha_\nu(t_c(\nu))}.$$

Since $\lim_{\nu\downarrow 0} f(u_\alpha(\nu),\nu)=0$ and $\lim_{\nu\uparrow m(u)} f(u_\alpha(\nu),\nu)=\infty$,

$$\lim_{\nu\downarrow 0}\frac{1}{\|v_\alpha(\nu)\|}v_\alpha(\nu)=\left(\frac{\partial}{\partial r}\right)_{\alpha_0(t_c(0))} \tag{29}$$

and

$$\lim_{\nu\uparrow m(u)}\frac{1}{\|v_\alpha(\nu)\|}v_\alpha(\nu)=\frac{1}{m(\alpha_{m(u)}(t_c(m(u))))}\left(\frac{\partial}{\partial\theta}\right)_{\alpha_{m(u)}(t_c(m(u)))}, \tag{30}$$

where $\|v_\alpha(\nu)\| := \sqrt{g(v_\alpha(\nu), v_\alpha(\nu))}$. Hence the subarc of the conjugate locus given by (25) is tangent to the parallel $r = r(\alpha_{m(0)}(t_c(0)))$ at $\alpha_{m(0)}(t_c(0))$ and to the opposite half meridian $\theta^{-1}(\pi)$ at $\alpha_{m(u)}(t_c(m(u)))$.

Next, we will argue about the first conjugate locus of q along β_ν, $\nu \in (0, m(u))$. It follows from [19, Proposition 7.2.3] that $\beta_\nu(t_f(\nu))$ is the first conjugate point of q along β_ν if and only if

$$\frac{\partial \theta_\beta}{\partial \nu}\bigl(r\bigl(\beta_\nu\bigl(t_f(\nu)\bigr)\bigr), \nu\bigr) = 0, \tag{31}$$

where

$$\theta_\beta(r, \nu) = \int_u^{2a-\xi(\nu)} f(r, \nu)\, dr + \int_r^{2a-\xi(\nu)} f(r, \nu)\, dr. \tag{32}$$

From (7), it is clear that

$$\theta_\beta(r, \nu) = \varphi(\nu) + \int_r^{2a-u} f(r, \nu)\, dr \tag{33}$$

holds. Hence, we have

$$\varphi'(\nu) + \int_{u_\beta(\nu)}^{2a-u} f_\nu(r, \nu)\, dr = 0, \tag{34}$$

where $u_\beta(\nu) := r(\beta_\nu(t_f(\nu)))$. By using the same argument as above, we may conclude that $u'_\beta(\nu) > 0$ on $(0, m(u))$. In particular, there is no cusp on the open arc defined by

$$\theta = \theta_\beta\bigl(u_\beta(\nu), \nu\bigr), \quad r = u_\beta(\nu), \tag{35}$$

on $\nu \in (0, m(u))$. It is easy to prove that

$$\lim_{\nu \downarrow 0} \frac{1}{\|v_\beta(\nu)\|} v_\beta(\nu) = \left(\frac{\partial}{\partial r}\right)_{\beta_0(t_f(0))} \tag{36}$$

and

$$\lim_{\nu \uparrow m(u)} \frac{1}{\|v_\beta(\nu)\|} v_\beta(\nu) = \frac{1}{m(\beta_{m(u)}(t_f(m(u))))} \left(\frac{\partial}{\partial \theta}\right)_{\beta_{m(u)}(t_f(m(u)))}, \tag{37}$$

where $v_\beta(\nu)$ denotes the velocity vector of the curve defined by (35). Since $\beta_{m(u)} = \alpha_{m(u)}$, by (30) and (37), we get

$$\lim_{\nu \uparrow m(u)} \frac{1}{\|v_\beta(\nu)\|} v_\beta(\nu) = \lim_{\nu \uparrow m(u)} \frac{1}{\|v_\alpha(\nu)\|} v_\alpha(\nu). \tag{38}$$

Therefore, the point $\alpha_{m(u)}(t_c(m(u))) = \beta_{m(u)}(t_f(m(u)))$ is a cusp of the first conjugate locus of q. Since M has a reflective symmetry with respect to the meridian passing through q, the points $\alpha_{m(0)}(t_c(m(u)))$ and $\beta_{m(0)}(t_c(m(u)))$ lying on the opposite half meridian to q are cusps of the conjugate locus. Therefore, the conjugate locus of q has exactly four cusps which consist of one pair lying on the parallel $r = 2a - r(q)$, the other lying on the opposite half meridian to q. $\square$

The following dual to Theorem 3.6 is also true, but we do not know examples satisfying the assumption in the theorem. We can at least say that numerical experiments very strongly suggest that such examples exist, e.g. $m(r) = \sin r - (1/100)\sin^3 r + (1/500)\sin^5 r$.

Theorem 3.7. *Let $(M, dr^2 + m(r)^2\, d\theta^2)$ denote a 2-sphere of revolution, where $m : (0, 2a) \to (0, \infty)$ is a smooth function satisfying (7). Assume that $\varphi : (0, m(a)) \to \mathbb{R}$ is well-defined, i.e., $m' \neq 0$ on $(0, a)$, and the half period function φ is monotone non-decreasing on $(0, m(a))$. If $\varphi''(\nu) \leqslant 0$ on $(0, m(a))$ and $\varphi''(\nu) < 0$ whenever $\varphi'(\nu) = 0$, then the first conjugate locus of any point q which is not a pole of M has exactly four cusps and the cut locus of q is a subarc of the opposite half meridian to q.*

Proof. It is proved in Theorem 3.6 that the first conjugate point of q along α_ν, β_ν, $\nu \in (0, m(u))$, is given by

$$\theta = \theta_\alpha\big(u_\alpha(\nu), \nu\big), \quad r = u_\alpha(\nu),$$

and

$$\theta = \theta_\beta\big(u_\beta(\nu), \nu\big), \quad r = u_\beta(\nu),$$

respectively. By making use of the assumption that $\varphi'(\nu) \geqslant 0$, we may prove that

$$u'_\alpha(\nu) > 0 \quad \text{and} \quad u'_\beta(\nu) < 0 \tag{39}$$

on $(0, m(u))$. The first claim is now clear from the argument in the proof of Theorem 3.6.

Since

$$\frac{\partial\theta_\alpha}{\partial\nu}\big(u_\alpha(\nu), \nu\big) = 0 \quad \text{and} \quad \frac{\partial\theta_\beta}{\partial\nu}\big(u_\beta(\nu), \nu\big) = 0,$$

we get

$$\frac{\partial}{\partial\nu}\theta_\alpha\big(u_\alpha(\nu), \nu\big) = u'_\alpha(\nu) f\big(u_\alpha(\nu), \nu\big), \qquad \frac{\partial}{\partial\nu}\theta_\beta\big(u_\beta(\nu), \nu\big) = -u'_\beta(\nu) f\big(u_\beta(\nu), \nu\big). \tag{40}$$

Hence, by (39), both functions $\theta_\alpha(u_\alpha(\nu), \nu)$ and $\theta_\beta(u_\beta(\nu), \nu)$ are strictly monotone increasing. In particular,

$$\theta_\alpha\big(u_\alpha(\nu), \nu\big) > \lim_{\nu\downarrow 0}\theta_\alpha\big(u_\alpha(\nu), \nu\big), \qquad \theta_\beta\big(u_\beta(\nu), \nu\big) > \lim_{\nu\downarrow 0}\theta_\beta\big(u_\beta(\nu), \nu\big) \tag{41}$$

for any $\nu \in (0, m(u)]$. Since the first conjugate points of $\alpha_{m(0)}$ and $\beta_{m(0)}$ lie in $\theta^{-1}(\pi)$ respectively,

$$\lim_{\nu\downarrow 0}\theta_\alpha\big(u_\alpha(\nu), \nu\big) = \lim_{\nu\downarrow 0}\theta_\beta\big(u_\beta(\nu), \nu\big) = \pi.$$

Here, we should recall that the geodesic polar coordinates (r, θ) are chosen so as to satisfy $\theta(q) = \pi$. Therefore, by (41), there is no conjugate point of q in $M \setminus \theta^{-1}(\pi)$. This implies that the cut locus of q is a subarc of the opposite half meridian $\theta^{-1}(\pi)$. □

4. Applications

Next, we apply our results to the problems introduced in Section 2.

Let g_λ be the family of analytic metrics on $\mathbb{S}^2$ defined by

$$g_\lambda = dr^2 + m_\lambda^2(r)\, d\theta^2$$

with

$$m_\lambda(r) = \sqrt{\lambda + 1}\, \sin r / \sqrt{1 + \lambda \cos^2 r}, \quad \lambda \geqslant 0. \tag{42}$$

It is clear that m_λ satisfies

$$m_\lambda(r) = m_\lambda(\pi - r),$$

so the Riemannian manifold $M_\lambda := (\mathbb{S}^2, g_\lambda)$ is a 2-sphere of revolution that is reflectively symmetric with respect to the equator $r = \pi/2$. This family of metrics contains the metric $\bar{g}$ associated with the averaged controlled Kepler equation introduced in Section 2 (set $\lambda = 4$), and defines a path between the following two remarkable metrics:

The case $\lambda = 0$ corresponds to the standard metric $g_0 = dr^2 + \sin^2 r\, d\theta^2$ on $\mathbb{S}^2$ with Gaussian curvature equal to unity, which is obtained by restricting the Euclidean metric on $\mathbb{R}^3$ to the sphere.

The case $\lambda = \infty$. In the limit case, the metric becomes

$$g_\infty = dr^2 + \tan^2 r\, d\theta^2$$

and is singular along the equator $r = \pi/2$. The Gaussian curvature is $-2/\cos^2 r$, which is strictly negative on each hemisphere and tends to $-\infty$ when r tends to $\pi/2$.

Lemma 4.1. *The Gaussian curvature G_λ of M_λ is given by*

$$G_\lambda(q) = \frac{(\lambda+1)(1-2\lambda\cos^2 r(q))}{(1+\lambda\cos^2 r(q))^2} \tag{43}$$

at a point $q \in M_\lambda$. Furthermore, we have

$$\left(\frac{\partial}{\partial r}\right)_q G_\lambda = \frac{2\lambda(\lambda+1)\sin 2r(q)}{(1+\lambda\cos^2 r(q))^3}\bigl(2-\lambda\cos^2 r(q)\bigr). \tag{44}$$

In particular, if $\lambda > 2$, then G_λ is not monotone along a meridian from a pole to the point on the equator.

Proof. By (42), we have

$$m'_\lambda(r) = \frac{(\lambda+1)\cos r}{(1+\lambda\cos^2 r)\sin r} m_\lambda(r). \tag{45}$$

Since

$$G_\lambda(q) = -\frac{m''_\lambda(r(q))}{m_\lambda(r(q))}$$

for $q \in r^{-1}(0, 2\pi)$, it follows from (45) that

$$G_\lambda(q) = \frac{(\lambda+1)(1-2\lambda\cos^2 r(q))}{(1+\lambda\cos^2 r(q))^2}. \tag{46}$$

By making use of (46), it is easy to show (44). □

Lemma 4.2. *Let a, b, c be positive numbers satisfying $c > b$. Then,*

$$\int \frac{1}{x(x+a)\sqrt{(x-b)(c-x)}}\,dx = \frac{2}{a}\left\{\frac{1}{\sqrt{bc}}\arctan\left(\sqrt{\frac{c}{b}}\,t\right) - \frac{1}{\sqrt{(a+c)(a+b)}}\arctan\left(\sqrt{\frac{a+c}{a+b}}\,t\right)\right\} \tag{47}$$

holds, where

$$t = \sqrt{\frac{x-b}{c-x}}.$$

In particular,

$$\int_b^c \frac{1}{x(x+a)\sqrt{(x-b)(c-x)}}\,dx = \frac{\pi}{a}\left\{\frac{1}{\sqrt{bc}} - \frac{1}{\sqrt{(a+c)(a+b)}}\right\}$$

holds.

Proof. From direct computation, we get

$$\frac{d}{dx}\left\{\frac{1}{\sqrt{bc}}\arctan\left(\sqrt{\frac{c}{b}}\,t\right) - \frac{1}{\sqrt{(a+c)(a+b)}}\arctan\left(\sqrt{\frac{a+c}{a+b}}\,t\right)\right\} = \frac{a}{2}\cdot\frac{1}{x(x+a)\sqrt{(x-b)(c-x)}}.$$

Hence, we obtain (47). □

Proposition 4.3. *For the 2-sphere of revolution M_λ, we get*

$$\varphi(\nu) = \pi - \frac{\lambda\pi\nu}{\sqrt{\lambda+1}\sqrt{\lambda+1+\lambda\nu^2}} \tag{48}$$

for each $\nu \in [0, m_\lambda(\pi/2)]$.

Proof. By putting $x = m_\lambda(r)^2$, we have, from (42) and (45),

$$dr = \frac{(1 + \lambda \cos^2 r) \tan r}{2(\lambda + 1)x} \, dx. \tag{49}$$

Since

$$x = m_\lambda(r)^2 = \frac{(\lambda + 1)(1 - \cos^2 r)}{1 + \lambda \cos^2 r}, \tag{50}$$

we obtain

$$\cos^2 r = \frac{\lambda + 1 - x}{\lambda x + \lambda + 1}. \tag{51}$$

Since

$$\tan^2 r = \frac{1}{\cos^2 r} - 1,$$

we have

$$\tan^2 r = \frac{(\lambda + 1)x}{\lambda + 1 - x}. \tag{52}$$

Combining (49), (51), and (52), we obtain

$$dr = \frac{(\lambda + 1)^{\frac{3}{2}}}{2(\lambda x + \lambda + 1)\sqrt{x(\lambda + 1 - x)}} \, dx. \tag{53}$$

Hence, we get

$$\varphi(\nu) = (\lambda + 1)^{\frac{3}{2}} \nu \int_{\nu^2}^{\lambda+1} \frac{1}{x(\lambda x + \lambda + 1)\sqrt{(x - \nu^2)(\lambda + 1 - x)}} \, dx. \tag{54}$$

It follows from Lemma 4.2 that

$$\varphi(\nu) = \pi - \frac{\lambda \pi \nu}{\sqrt{\lambda + 1}\sqrt{\lambda + 1 + \lambda \nu^2}}. \qquad \square$$

Theorem 4.4. *If $\lambda > 0$, then, for each point q of M_λ distinct from a pole, the cut locus of q is a subarc of the antipodal parallel to q and the first conjugate locus of q has exactly four cusps.*

Proof. It is clear that

$$\varphi'(\nu) = \frac{-\lambda \pi \sqrt{\lambda + 1}}{(\lambda + 1 + \lambda \nu^2)^{\frac{3}{2}}}$$

and

$$\varphi''(\nu) = \frac{3\pi \lambda^2 \nu \sqrt{\lambda + 1}}{(\lambda + 1 + \lambda \nu^2)^{\frac{5}{2}}}.$$

In particular,

$$\varphi'(\nu) < 0 < \varphi''(\nu)$$

on $(0, m_\lambda(\pi/2))$, if $\lambda > 0$. The claims of Theorem 4.4 are now clear from Lemma 3.3, Theorems 3.5 and 3.6. $\square$

Remark 1. If $\lambda > 2$, then the Gaussian curvature of M_λ is not monotone along a meridian from a pole to the point on the equator. Therefore, the family M_λ, $\lambda > 2$, is a new example which has the simple cut locus structure.

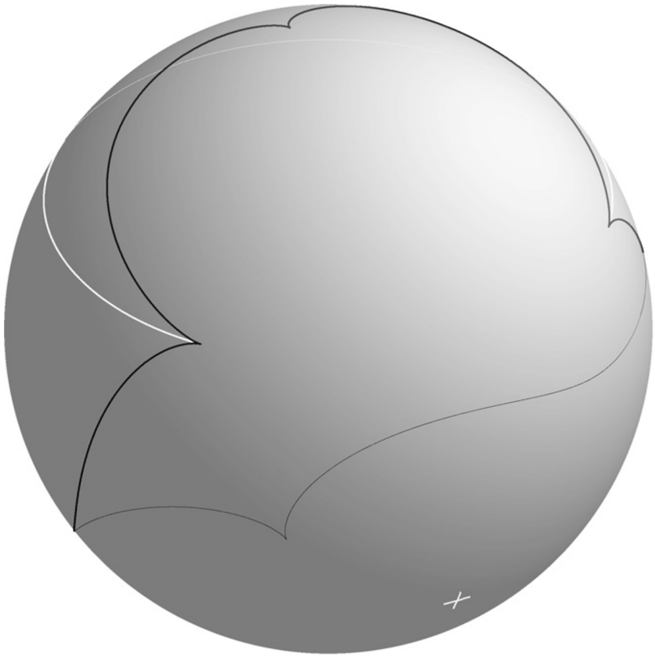

Fig. 1. Conjugate locus (black) and cut locus (white) in averaged orbital transfer ($\lambda = 4$) for an initial condition, indicated by a white cross, away from the poles.

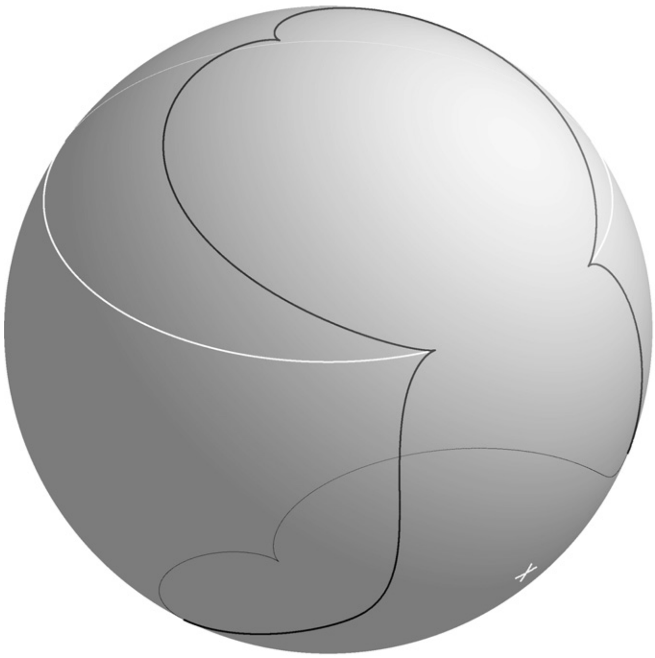

Fig. 2. Conjugate locus (black) and cut locus (white) in the singular case ($\lambda = \infty$) for an initial condition, indicated by a white cross, not on the equator and away from the poles.

In Fig. 1 we present conjugate and cut loci in averaged orbital transfer ($\lambda = 4$). Since to lose optimality an extremal trajectory has to cross the equator, $e = 1$, we conclude that extremals are optimal in the physical elliptic domain. The conjugate and cut loci for $\lambda = \infty$ are given in Figs. 2 and 3. In this singular case, since the curvature outside the equator is strictly negative, geodesics starting from a point not on the equator have to cross it to have conjugate points,

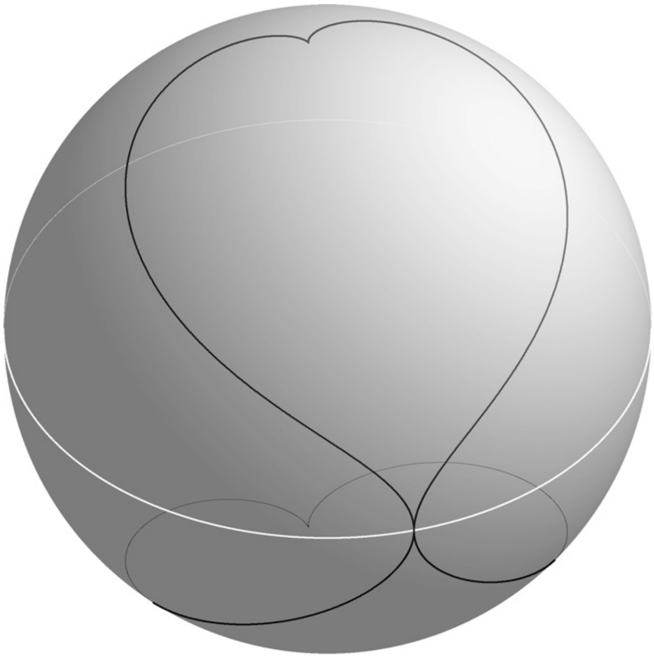

Fig. 3. Conjugate and cut loci in the singular case ($\lambda = \infty$) for an initial condition on the equator.

the cut locus still being included in the antipodal parallel. For a point on the equator, the conjugate and cut loci accumulate near the point itself, and the cut locus is the equator minus the point. The first two cusps of the conjugate locus disappear and are replaced by a contact of order two at the initial point [1,8].

5. The singular case

In this section, we outline the analysis of the case where the metric $g = dr^2 + m^2(r)\,d\theta^2$ on the 2-sphere of revolution is singular on the equator and we refer to [8] for more details. The singularity encountered coming from the Grusin model example was analyzed by [1] and is described near the point identified to 0 by the local model:

$$g_s = dx^2 + \frac{dy^2}{x^2}$$

and the analysis can be extended to the case of order p:

$$g_s = dx^2 + p^2\frac{dy^2}{x^{2p}}$$

called a *generalized Grusin singularity*.

While the Riemannian metric and the Gaussian curvature explode when approaching the y-axis, still the extremal curves are still described by a smooth Hamiltonian system

$$H = \frac{1}{2}\left(p_x^2 + \frac{x^{2p}}{p^2}p_y^2\right).$$

They are associated to SR-geometry in dimension 3.

Indeed, the case $p = 1$ is deduced from the Heisenberg case with corresponding Hamiltonian

$$H = \frac{1}{2}\left[\left(p_x^2 + p_y^2\right) - 2p_z(xp_y - yp_x) + (x^2 + y^2)p_z^2\right]$$

since using cylindrical coordinates

$$H = \frac{1}{2}\left(p_r^2 + (p_\theta/r - rp_z)^2\right).$$

As θ is cyclic, p_θ is a first integral and for $p_\theta = 0$, the reduced Hamiltonian in the (r, z) space has the desired singularity:

$$H = \frac{1}{2}\left(p_r^2 + r^2 p_z^2\right).$$

Similarly, the case of order 2 can be deduced from the so-called Martinet flat case, with Hamiltonian

$$H = \frac{1}{2}\left(\left(p_x + \frac{y^2}{2} p_z\right)^2 + p_y^2\right).$$

As x is cyclic, p_x is a first integral and for $p_x = 0$, the reduced Hamiltonian is of the form

$$H = \frac{1}{2}\left(p_y^2 + \frac{y^4}{4} p_z^2\right).$$

In particular in both cases $p = 1, 2$ the extremal flow, the conjugate and cut loci can be deduced from the analysis of the SR-problem, see [9].

- *The case* $p = 1$: The extremal trajectories with initial condition $x(0) = y(0) = 0$ and parameterized by arc-length are given, with $\kappa = p_y(0) \geqslant 0$, by:
  - $\kappa = 0$: $x(t) = \pm t$, $y(t) = 0$,
  - $\kappa > 0$: $x(t) = \pm(\sin \kappa t)/\kappa$, $y(t) = t/(2\kappa) - (\sin 2\kappa t)/(4\kappa^2)$

  while extremals for $\kappa < 0$ are obtained by reflection with respect to the x-axis.
  For $\kappa > 0$, the first conjugate time is at $t_{1c} = \tau/\kappa$, $\tau \simeq 4.5$, while, due to symmetry, optimality is lost at time π/κ, when crossing the y-axis.
- *The case* $p = 2$: The extremal trajectories with initial condition $x(0) = y(0) = 0$ and parameterized by arc-length, with $\kappa = p_y(0) \geqslant 0$ are:
  - $\kappa = 0$: $x(t) = t$, $y(t) = 0$,
  - $\kappa > 0$: $x(t) = -(2k/\sqrt{\kappa}) \operatorname{cn} u$, $y(t) = (2/(3\kappa^{3/2}))[(2k^2 - 1)(E(u) - E(K)) + k'^2 t\sqrt{\kappa} + 2k^2 \sin u \operatorname{cn} u \operatorname{dn} u]$,

  where $u = K + t\sqrt{\kappa}$, $k^2 = k'^2 = 1/2$, and the curves deduced from the previous ones using the reflections with respect to the x and y-axis.
  For $\kappa > 0$, the first conjugate time is at time $t_{1c} \simeq 3K/\sqrt{\kappa}$, while due to symmetries optimality is lost at time $2K/\sqrt{\kappa}$, when crossing the y-axis.

Hence, for both cases, we have the same geometric situation, optimality is lost due to the symmetry with respect to the y-axis and the conjugate and cut loci are disjoint, because the first conjugate point occurs after the crossing of the y-axis. It can be generalized to any order using quasi-homogeneity [8].

Proposition 5.1. *Consider a metric of the form $g_s = dx^2 + p^2\,dy^2/x^{2p}$. Then at the origin the conjugate and cut loci are disjoint, the cut locus is the y-axis minus* 0*, while the conjugate locus is a set of the form $y = \pm c_p x^{p+1}$ minus* 0.

Hence, the model gives locally the cut and conjugate loci of a point of the equator observed in Fig. 3, which corresponds to a singularity of order 1.

Having resolved the singularity at the equator, we can extend the result of the regular case.

Definition 5.2. Consider a metric on a 2-sphere of revolution of the form $dr^2 + m^2(r)\,d\theta^2$, with m' non-zero on $(0, \pi/2)$ and $m(\pi - r) = m(r)$, smooth everywhere but the equator for which we have a Grusin singularity. The extremal flow is called *tame* if the half period function $\varphi(\nu)$ of Section 3 is monotone non-increasing for $\nu \in]0, m(\pi/2) = +\infty[$.

Since the extremal flow remains smooth, the analysis of the regular case can be extended.

Proposition 5.3. *In the tame case, we have*

1. *The cut locus of a point different from a pole and not on the equator is a subset of the antipodal parallel.*
2. *The cut locus of a point on the equator is a subset of the equator accumulating at the point.*

6. Conclusion

The contribution of this article is twofold. First of all we give a simple and computable criterion to decide whether the cut and conjugate loci on a surface of revolution are the simplest possible. It is based on two case studies in optimal control, in the framework of Hamiltonian dynamics. Such specialized results are important because even on a surface of revolution the computation of these objects is in general intractable. This provides a bridge between Riemannian geometry and optimal control, with promising extensions to the singular case and the Zermelo navigation problem on Riemannian manifolds [2]. Secondly, we give a neat proof of the structure of the conjugate and cut loci in coplanar orbital transfer, which were previously computed in [4] using the explicit parameterization of the extremal flow for each initial point where it is sufficient to consider the half period mapping. In the tangential case, they were obtained using a normal form and numerical simulations. Applications and generalizations of these results, using Hamiltonian formalism, will provide very significant improvements in the understanding of Lindblad equations describing the interaction of a two-level quantum system controlled by a laser with an environment. Based on the theoretical concept and results of our analysis in the Riemannian case, this will be analyzed in a forthcoming article. Roughly speaking, if the interaction is weak, it defines a Zermelo navigation problem on a two-sphere of revolution for which conjugate and cut loci have a similar simple structure.

References

[1] A. Agrachev, U. Boscain, M. Sigalotti, A Gauss–Bonnet like formula on two-dimensional almost Riemannian manifolds, Preprint SISSA 55/2006/M, 2006.
[2] D. Bao, C. Robles, Z. Shen, Zermelo navigation on Riemannian manifolds, J. Differential Geom. 66 (3) (2004) 377–435.
[3] M. Berger, A Panoramic View of Riemannian Geometry, Springer-Verlag, Berlin, 2003.
[4] B. Bonnard, J.-B. Caillau, Optimality results in orbit transfer, C. R. Acad. Sci. Paris, Ser. I 345 (2007) 319–324.
[5] B. Bonnard, J.-B. Caillau, R. Dujol, Energy minimization of single-input orbit transfer by averaging and continuation, Bull. Sci. Math. 130 (8) (2006) 707–719.
[6] B. Bonnard, J.-B. Caillau, Riemannian metric of the averaged energy minimization problem in orbital transfer with low thrust, Ann. Inst. H. Poincaré Anal. Non Linéaire 24 (3) (2007) 395–411.
[7] B. Bonnard, D. Sugny, Optimal control with applications in space and quantum dynamics, Preprint Institut math. Bourgogne, 2007.
[8] B. Bonnard, J.B. Caillau, M. Tanaka, One-parameter family of Clairaut–Liouville metrics with application to optimal control, HAL preprint No 00177686 (2007), url: hal.archives-ouvertes.fr/hal-00177686.
[9] B. Bonnard, M. Chyba, Singular Trajectories and Their Role in Control Theory, Springer-Verlag, Berlin, 2003.
[10] H. Gluck, D. Singer, Scattering of geodesic fields I and II, Ann. of Math. 108 (1978) 347–372, and 110 (1979) 205–225.
[11] J. Gravesen, S. Markvorsen, R. Sinclair, M. Tanaka, The cut locus of a torus of revolution, Asian J. Math. 9 (2005) 103–120.
[12] J.J. Hebda, Metric structure of cut loci in surfaces and Ambrose's problem, J. Differential Geom. 40 (1994) 621–642.
[13] J.I. Itoh, K. Kiyohara, The cut loci and the conjugate loci on ellipsoids, Manuscripta Math. 114 (2004) 247–264.
[14] K. Fleischmann, Die geodätischen Linien auf Rotationsflächen, Liegnitz, Dissertation, Seyffarth, 1915. Persistent URL: http://www-gdz.sub.uni-goettingen.de/cgi-bin/digbib.cgi?PPN317487884.
[15] S.B. Myers, Connections between Differential Geometry and Topology: I. Simply connected surfaces and II. Closed surfaces, Duke Math. J. 1 (1935) 376–391, and 2 (1936) 95–102.
[16] H. Poincaré, Sur les lignes géodésiques des surfaces convexes, Trans. Amer. Math. Soc. 6 (1905) 237–274.
[17] T. Sakai, Riemannian Geometry, Transl. Math. Monogr., vol. 149, American Mathematical Society, Providence, RI, 1996.
[18] K. Shiohama, M. Tanaka, Cut loci and distance spheres on Alexandrov surfaces, in: Actes de la table ronde de géométrie différentielle en l'honneur de Marcel Berger, in: Collection Soc. Math. France, vol. 1, Séminaires & Congrès, 1996, pp. 531–560.
[19] K. Shiohama, T. Shioya, M. Tanaka, The Geometry of Total Curvature on Complete Open Surfaces, Cambridge Tracts in Mathematics, vol. 159, 2003.
[20] R. Sinclair, M. Tanaka, A bound on the number of endpoints of the cut locus, LMS J. Comput. Math. 9 (2006) 21–39.
[21] R. Sinclair, M. Tanaka, The cut locus of a two-sphere of revolution and Toponogov's comparison theorem, Tohoku Math. J. 59 (2) (2007) 379–399.
[22] D. Sugny, C. Kontz, H. Jauslin, Time-optimal control of a two-level dissipative quantum system, Phys. Rev. A 76 (2007) 023419.